

192

NASA CONTRACTOR REPORT



NASA CR-128

NASA CR-128

N 64 32832

ACCESSION NUMBER

SPACE

(THRU)

CODE

CATEGORY

RELIABILITY AND REDUNDANT CIRCUITRY

by P. R. Dennis and Sundarum Seshu

Prepared under Contract No. NASw-778 by

CLYDE WILLIAMS AND COMPANY

Columbus, Ohio

for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • OCTOBER 1964

RELIABILITY AND REDUNDANT CIRCUITRY

By P. R. Dennis and Sundarum Seshu

Distribution of this report is provided in the interest of information exchange. Responsibility for the contents resides in the author or organization that prepared it.

Prepared under Contract No. NASw-778 by
CLYDE WILLIAMS AND COMPANY
Columbus, Ohio

for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Office of Technical Services, Department of Commerce,
Washington, D.C. 20230 -- Price \$1.00

SUMMARY

32832

Within the complex field of reliability, there are many technical disciplines which are large and complicated efforts in themselves. These subfields of reliability encompass virtually all technologies and touch every phase of product development from the initial design concept through final product production and ultimate product use. This report "Reliability and Redundant Circuitry" is concerned with only one area within the overall field of reliability, viz., the use of redundancy techniques to improve the reliability of systems.

One way to make a system more reliable is to improve the reliability of each of its parts. Unfortunately, in many cases product reliability has been pushed so close to its limits that further improvement may be uneconomical or even impossible. In cases of this type, it may be found necessary to design the system so that it functions properly even when some of its parts fail. To overcome malfunctions caused by failed parts, redundant (or extra) parts must be used in the system--parts which would be quite unnecessary if no failures ever occurred. Many interesting and complex theories have been built up around the basic redundancy concept. Some of these theories have direct practical application--some, as yet, do not.

The report is not a "textbook" on redundancy theory and applications, but rather pulls together diverse work in the field of redundancy involving both NASA and non-NASA contributions to indicate the state-of-the-art. The references used in the report and listed in its bibliography comprise an excellent cross section of work accomplished in the field of redundancy. They include both report and open literature for the 10-year period 1954-63. Individual references are discussed, their unique contributions are noted, and some comments are made on the validity and practical aspects of the individual works. The report also delineates various types of redundancy and the applications of each type.




TABLE OF CONTENTS

RELIABILITY INTENSIVELY STUDIED	1
DEFINITIONS	1
THE RELIABILITY FUNCTION	2
BASIC REDUNDANCY THEORY	3
TECHNIQUES FOR REDUNDANCY	6
ACTIVE REDUNDANCY	7
Parallel Redundancy	7
Triple Modular Redundancy	9
Quad Redundancy	10
UNMAINTAINED REDUNDANT SYSTEMS	12
REDUNDANCY WITH MAINTENANCE	16
SELF REPAIRING REDUNDANT SYSTEMS	18
COMMERCIAL APPLICATIONS	19
ACHIEVEMENT OF RELIABILITY	24
REDUNDANCY NOT A CURE-ALL	24

RELIABILITY AND REDUNDANT CIRCUITRY

By P. R. Dennis and Dr. Sundarum Seshu*

Reliability Intensively Studied

The reliability of electronic equipment became a subject of intensive study shortly after World War II, largely at the insistence of the armed services. The literature in this general field began to mushroom about 10 years ago, and currently includes thousands of papers and reports.

This report summarizes only a small portion of the field, namely, the use of redundancy to improve the reliability of digital systems. Its purpose is to indicate the state of the art in general theory and to comment on some specific applications.

In the field of reliability alone, one can list 500 or more publications, but the majority of these are either tutorial or repetitive. Only major contributions are included in the highly selective bibliography.

Definitions

Accepted definitions of basic concepts in reliability theory follow.

The reliability of a particular piece of equipment, denoted by the function $R(t)$, is the a priori probability that the equipment will perform its function properly at any instant up to the time t .

In this definition, the time origin ($t = 0$) is assumed to be the instant at which the equipment is put into service. Since $R(t)$ is an a priori probability, it follows that $R(t)$ is a monotonically decreasing function of t . Also, $R(0)$ is taken to be 1. The a priori probability of failure is $P(t) = 1 - R(t)$.

The mean life (sometimes also referred to as mean time to failure or mean time between failures depending on the mathematical model) is that a priori lifetime without failure which is expected of the equipment.

*Dr. Seshu is a consultant for CLYDE WILLIAMS AND COMPANY and is a full time staff member of the Department of Electrical Engineering, University of Illinois, Urbana, Illinois.

Thus, if M stands for mean life (see, for example, Reference 33)

$$(1) \quad M = \int_0^{\infty} R(t) dt$$

For systems which are not maintained, and which do not have definite periods of usefulness, mean life is a more meaningful characteristic than reliability.

For systems which are maintained the characteristic of availability is useful.

Availability (\bar{A}) is a ratio which gives a measure of the portion of the time during which a piece of equipment functions usefully.

$$(2) \quad \bar{A} = \frac{M}{M + T_R}$$

where M is the mean time between failures and T_R is the mean repair (or maintenance) time.

Much of this nomenclature becomes confusing when applied to redundant systems. It has been suggested that the term reliability be used only when all the components are known to be working at $t = 0$. If it is known that only the equipment or the total system is working at $t = 0$, it has been suggested that the term continuance be used instead. However, the term continuance has not been universally adopted.

In the study of reliability, one must also distinguish between catastrophic and intermittent failures. A component is said to fail catastrophically at $t = t_0$, if the probability $P(t)$ that the component will function properly is identically zero for $t > t_0$; otherwise the failure is intermittent.

We have taken the terms component, equipment, failure, etc., to be basically undefined. Operationally speaking, a component is the lowest replaceable part.

The Reliability Function

It is known that for complex, nonredundant electronic equipment consisting of a large number of components, the reliability, $R(t)$, is given by

$$(3) \quad R(t) = e^{-\lambda t}$$

where λ is the failure rate of the equipment (see, for example, Reference 12). The equipment reliability is independent of the reliability of individual components. One property of this Poisson function is that if the equipment is known to be working at t_0 , the (conditional) reliability (R_c) remains exponential:

$$(4) \quad R_c(t-t_0) = R(t-t_0) = e^{-\lambda(t-t_0)}$$

So far as physical systems are concerned, two assumptions are implicit in the above description. The first is that the equipment has gone through a "debugging" or "burn-in" period during which poorly manufactured components are replaced. The second is that the mean life of the equipment is very short compared with the mean life of the components so that the "wear-out" period is not encountered. That is, the encountered failures are of random causes.

One striking feature of the reliability function shown as Equation 3 is the rate at which it falls off. Within a relatively short time the reliability of complex electronic equipment becomes very low. At $t = M$, for example, the reliability is only 0.37. And at $t = 0.5M$, the reliability is only about 0.7.

The purpose of redundancy in a system is to appreciably improve the reliability of a complex piece of equipment during a period shorter than the mean life of the corresponding nonredundant system. This purpose can be fulfilled, as a comparison of the curves of $R(t)$ for nonredundant and redundant systems respectively will show. (Figure 1.)

Basic Redundancy Theory

One way to make a system more reliable is to improve the reliability of each of its parts. Unfortunately, in many cases component reliability has been pushed so close to its limits that further improvement may be uneconomical or even impossible. In cases of this type, it may be found necessary to design the system so that it functions properly even when some of its parts fail. To accomplish this redundant, or extra, parts must be used to overcome errors caused by the failed parts. These redundant parts would be quite unnecessary if no errors ever occurred.

Redundancy theory is closely related to the field of information theory. Work in either field can and often does supplement work in its sister field.

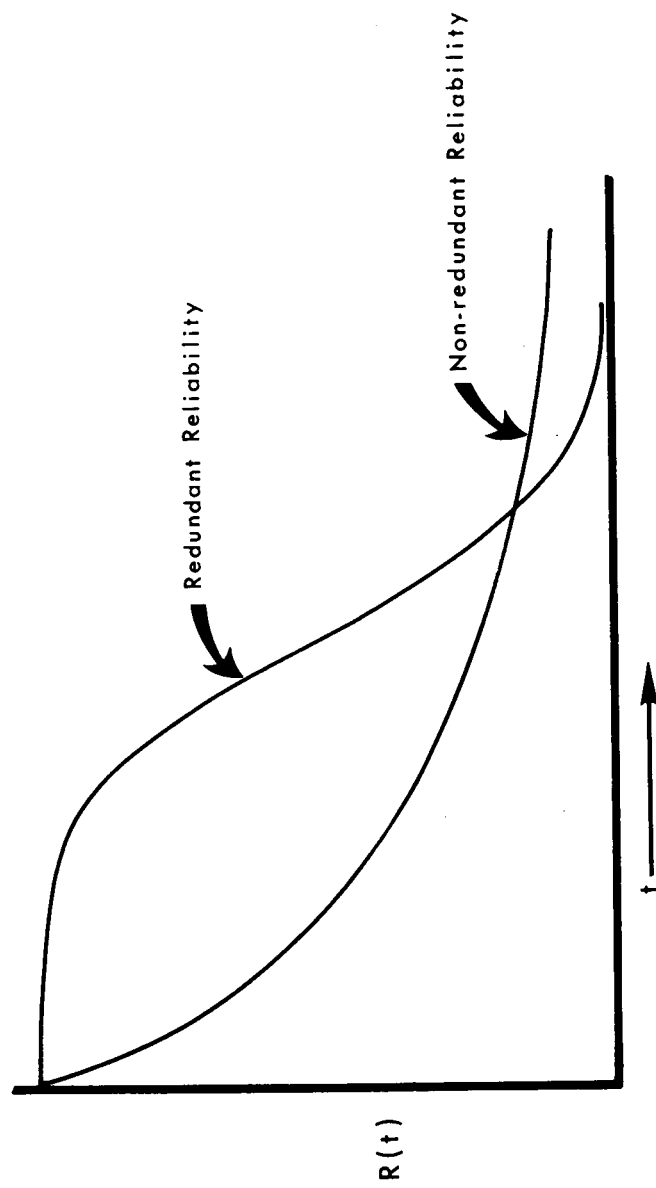


FIGURE 1. COMPARISON OF REDUNDANT AND NON-REDUNDANT RELIABILITY

However, where information theory deals with the handling of unreliable information with transmission methods that are assumed to be reliable, redundancy theory considers the handling of unreliable information with unreliable transmission methods.

Apart from work done at the Naval Research Laboratory (see Reference 13), most research on redundancy follows the pioneering contribution of von Neumann (54). Von Neumann's work has been generally accepted without criticism. In an attempt to set up an analog of the human (or animal) nervous system, he attributed to the receiving mechanism (the brain) certain capabilities which must be part of the logic in an inorganic system.

Briefly, von Neumann considered two types of systems. In the first system, each function to be performed is triplicated, and a majority decision element is added to determine the output. For n stages in a nonredundant circuit, n elements would suffice. However, the corresponding redundant configuration would require 3^n elements in order to avoid the downgrading of reliability in multistage logic. The majority decision element itself must have a reliability greater than 0.84. The reliability of the majority decision element sets the upper limit to the reliability of the circuit as a whole.

In the second system proposed by von Neumann, bundles of wires are used to carry signals. The signal itself serves as a 3-way reliability check according to the following convention: if fewer than a certain number (later fixed at 7 percent) of the wires signal 1, the actual signal is considered to be zero. If fewer than the same number of wires signal 0, the signal is considered to be 1. More than 7 percent and fewer than 93 percent of the wires signaling 0 or 1 is considered proof of malfunction.

This method of determining the true output demands that the equipment must incorporate a complicated (and therefore unreliable) mechanism, and this mechanism von Neumann put into the "brain." The reliability of this added mechanism again sets the upper limit to the reliability of the system as a whole. Little gain in reliability is obtainable with fewer than 20,000 wires. Thus, such provision for redundancy adds disproportionate bulk to equipment, and is impracticable with conventional present-day hardware.

Von Neumann's work centered around two basic electronic gates: the majority gate and the Sheffer stroke gate.

Implicit in his theory are the following assumptions:

1. All failures are intermittent.
2. Failures are statistically independent.
3. The system is stationary (independent of time).
4. Failures are independent of inputs.

The first and third assumptions are critical, but are often overlooked. The consequences of using redundancy in electronic digital systems where catastrophic failures are more common than intermittent failures are still not well understood. The third assumption implies that one is working to a time base which is an extremely small fraction of the mean life of a component. (Approximately $\frac{0.1}{N}$, where N is the number of components.)

Moore and Shannon (40) followed von Neumann in their investigation of redundancy applied to relay contact networks. Their assumptions were the same. Their results were, however, drastically different. In the first place, they did not have to assume that the relays were 84 percent reliable. A probability of failure of anything less than 50 percent was sufficient. (If $p > 0.5$, the contact is of the opposite type. If $p = 0.5$, the behavior is random.) In the second place, they obtained their redundant circuits simply by replacing each contact with a standard circuit. Finally, the redundancy required for a specific improvement in reliability was very small compared with that required for von Neumann's electronic cases (redundancy ratios of the order of 10 as against 10,000).

Attempts to apply the Moore-Shannon networks to electronic hardware have generally been unsuccessful or invalid. The closest practical circuit is the quad configuration mentioned later. An abstract formalization of their results has been given by Birnbaum et al (5).

Since the appearance of von Neumann's paper there have been numerous publications on the reliability of systems incorporating different configurations of redundant circuitry. The most comprehensive study to date is a paper by Flehinger (21) which is discussed later.

Techniques for Redundancy

It is appropriate at this point to classify different procedures utilizing redundancy.

Active, or Wired-in Redundancy is the simplest type of redundancy procedure. It was proposed by von Neumann and by Moore and Shannon. All components in a system are in operation at the same time, and contribute in some way to the output. This method is valuable for unmanned systems which are required to operate for relatively short periods with maximum protection against intermittent failures (only).

Passive or Standby Redundancy is commonly used in manned systems where cost of downtime is prohibitively high, as would be the case in a defense

installation. All the components in the system are active, but the output is taken from only one "chain" of components. When this chain fails, the output is taken from another chain. If the system is manned, the chain which has failed is then repaired. This arrangement offers the greatest increase in lifetime for a given redundancy and will be discussed at greater length later.

System Redundancy with Repair is a variation of Passive or Standby Redundancy. Several identical subsystems operate simultaneously and the output is by majority decision. When a subsystem output disagrees with the majority output, it is removed from service and repaired.

The main difference between Passive or Standby Redundancy and System Redundancy with Repair is the "level" at which redundancy is applied. But, the level of application has a significant effect on the reliability as well as on the feasibility of using redundancy. Passive or Standby Redundancy and System Redundancy with Repair are preferred for systems which are required to operate for long periods of time (in comparison with the mean life of the corresponding nonredundant system). In such applications, the probability of catastrophic failures should be minimized.

In the following sections, all of these procedures are considered in detail.

In the first section, only systems of the first type are discussed. Unmaintained systems incorporating Passive or Standby Redundancy and System Redundancy with Repair are considered in the second section, while maintained systems of the same types are considered in the third section. In the fourth section, self-redundant systems are discussed.

Carroll (7) has given a more detailed introductory survey of the various types of redundant systems.

Active Redundancy

Parallel Redundancy

The simplest type of redundancy is the simple parallel redundancy at the system level (see Figure 2). In this type, three or more identical (nonredundant) subsystems receive the input simultaneously. Their outputs are fed into a majority decision element M . The output of M is the value (1 or 0) of the output of the majority of the subsystems S_i . If the subsystems S_i have reliabilities r , and the majority element has reliability r_M , the system reliability can be

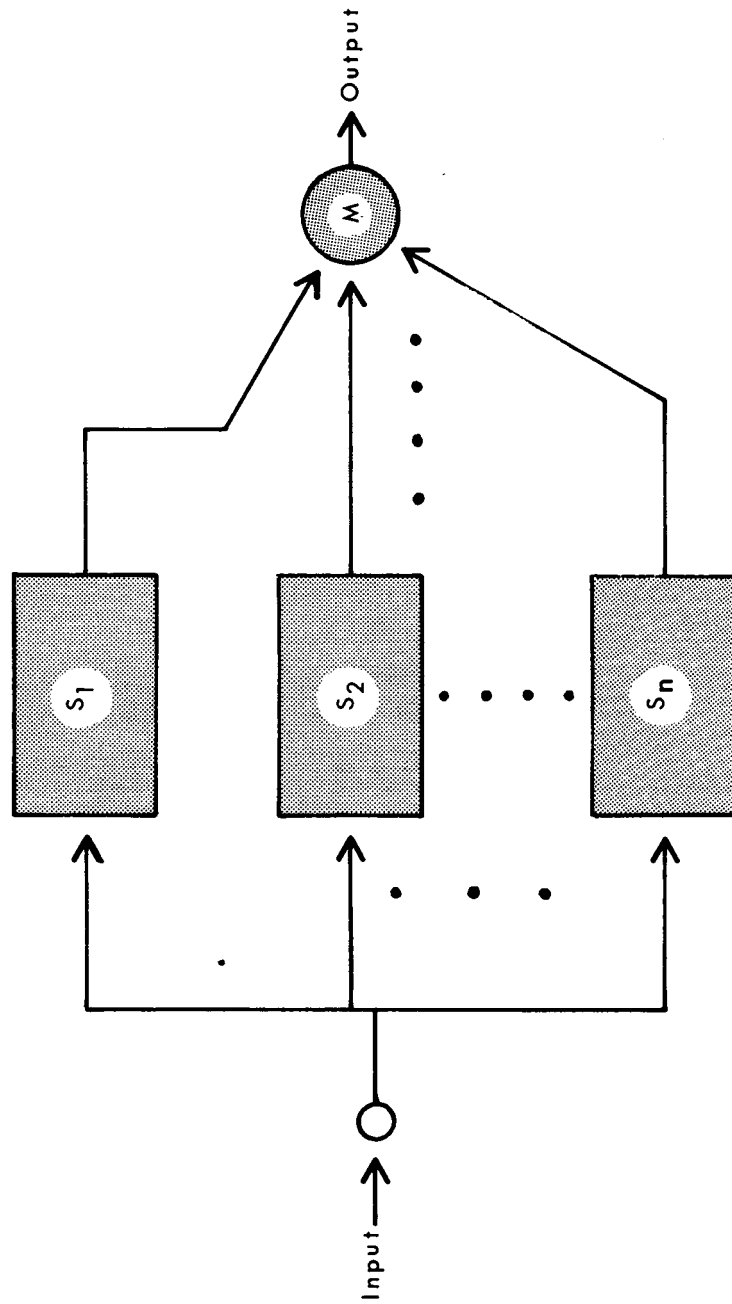


FIGURE 2. SIMPLE PARALLEL REDUNDANCY

calculated to be

$$(5) \quad R(t) = r_M \sum_{k=0}^n \binom{n}{k} r^k (1-r)^{n-k}$$

$$k = \left\lceil \frac{M}{2} \right\rceil$$

where $\left\lceil \frac{M}{2} \right\rceil$ is the smallest integer greater than $\frac{M}{2}$ and $\binom{n}{k}$ is the binomial coefficient. The curve of this equation has the general shape of the curve shown in Figure 1.

Using procedures similar to those of Kletsky (35), Equation 5 can be integrated to give the probable mean life of the system. Note that the mean life of the system increases as $\log n$ (where n is the number of parallel systems) and that it is limited by the mean life of the majority element.

Pierce (43, 44, 45) has suggested an interesting variation in which the output of each subsystem, S_i , is first passed through a variable gain element, a_i (made analog). After summation, the results are fed through a threshold device. A decision element sets up the gain a_i by comparing the actual system output and the subsystem output. Pierce claimed very large increases in reliability (of the order of 500). But he assumed that the adaptive gain a_i , the decision element which sets up a_i , the vote taker, and the threshold element are all failure free. This is clearly unrealistic. In addition, the distribution of reliabilities over the aggregate of the subsystems must be known. (He assumes that this follows a Gaussian curve.) Another assumption is that failures are intermittent. No assumption can be made that the decision element -- a very complicated device -- is more reliable than the system itself. Even with all these assumptions, the gain in mean life is only about three times. A discussion of the decision processes required was given by Farrell (19).

Price (47) attempted to extend the paralleling concept to catastrophic failures, but his formulae appear to be invalid (see discussion by Flehinger [24]).

Triple Modular Redundancy

Several authors have considered structures close to the von Neumann configuration, namely, the so-called triple modular redundancy (6, 38). The

digital system is divided into individual modules. Each module is then replaced by three modules, followed by three majority elements as shown in Figure 3. At the output end, there is a single majority element. If the voting circuits are assumed to be perfect, the (instantaneous) reliability can be made as close to 1 as desired by dividing the system into a sufficiently large number of modules. If R_0 is the reliability of a system of m modules in cascade so that the module reliability is $R_0^{1/m}$, the reliability of the triple redundant system is:

$$(6) \quad R = (3R_0^{2/m} - 2R_0^{3/m})^m$$

If, however, the voting circuits are not perfect, but have a reliability R_v , the reliability of the redundant system is:

$$(7) \quad R = (3R_v^2 R_0^{2/m} - 2R_v^3 R_0^{3/m})^m R_v$$

(The final multiplier R_v was omitted by Lyons and Vanderkulk [38]).

With extremely reliable voting circuits, say $R_v = 0.999$, it is possible to achieve reliabilities of the order of 0.95 for a value of t equal to the mean life of the nonredundant system. Increases in mean life are, however, only modest. Mean life of the triple modular system is about 1.5 times that of the nonredundant system.

The modular redundancy has to be applied at a level of about 1/60th the size of the system for these values to hold. That is, the system must be divided into at least 60 modules which are then made redundant.

For the simple majority decision system there is no measurable increase in mean life as illustrated by Dickinson and Walker (16). Thus, the lower level gain of $\frac{3}{2}$ should be considered good.

Quad Redundancy

Another type of redundancy which protects equipment against intermittent failures is the so-called "quad" redundancy. There are two conceptually different schemes which go by this term.

The first type of quad redundancy -- and the more common of the two --

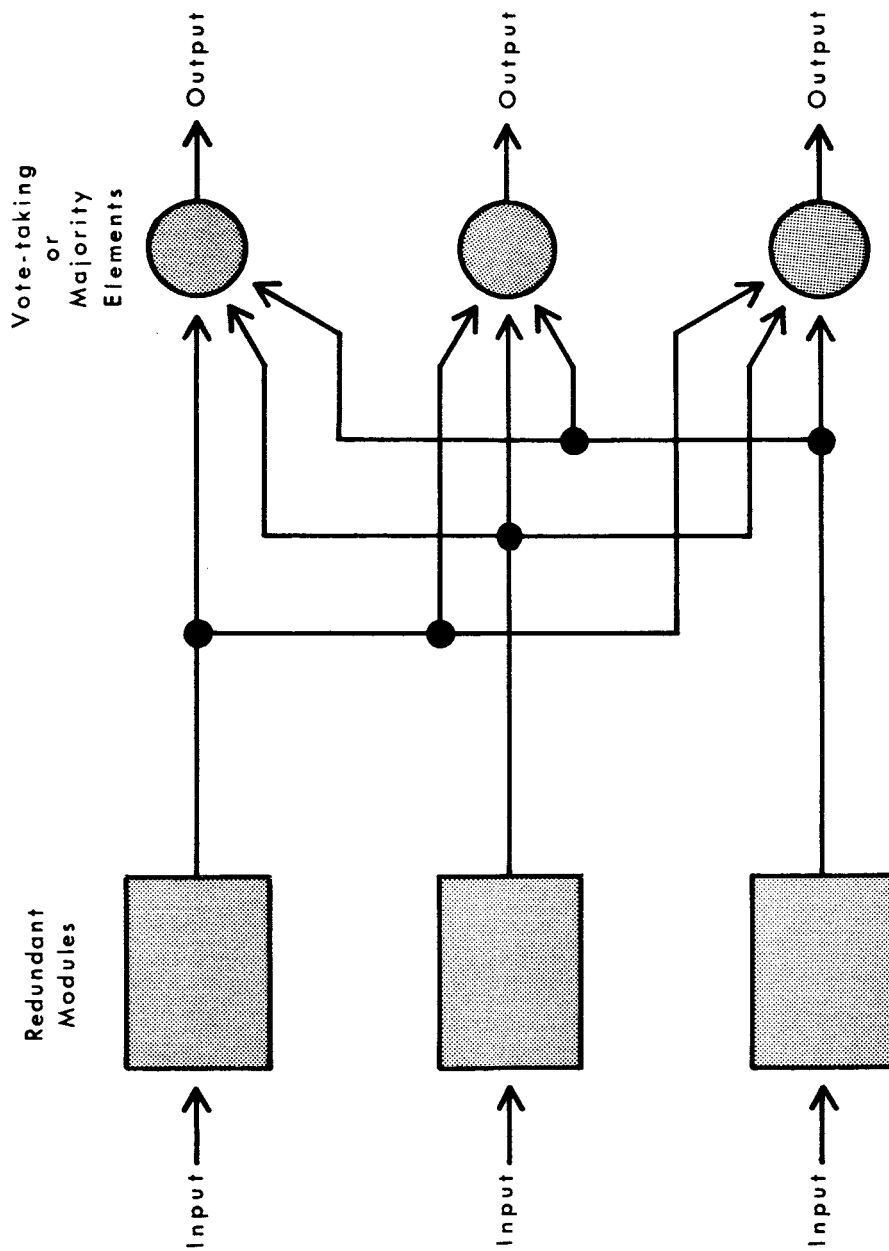


FIGURE 3. TRIPLE MODULAR REDUNDANCY

is applied to the configurations shown in Figure 4. The two arrangements (which are duals of each other) were originally given by Moore and Shannon (40) for relay contacts. They have since been used for diodes, resistors, and even complete transistor circuits. Their application to switching diodes is straightforward and the original formulae hold. Their applicability in other cases is questionable, since the analog constants (resistance, gain, loading, etc.) change when one module opens or shorts. In any case, protection is provided against intermittent open and short-circuit failures only. A simple discussion of this topic has been presented by Plait (46).

The second type of quad redundancy is an intriguing arrangement put forward by Tryon (53). Shown in Figure 5, it is applicable to monotone logic such as AND/OR. It offers full protection against any single failure and against pairs of failures separated by several stages of logic. Again, intermittent failures only are considered. Formulae for the reliability or mean life of the quadded logic are not established. Tryon also provides a detailed discussion of the application of quadded logic to various parts of a digital computer.

Unmaintained Redundant Systems

In this section the following redundant system is considered:

1. The system is unmanned.
2. Not all the available equipment is contributing to the output; some of the equipment is thus "spare."
3. The spares may consist of modules, or of complete subsystems which are capable of performing the system function, or of any intermediate level of operation.
4. There are switching arrangements to bring these spares into operation.
5. The spares may be either "active" (alive) or on the shelf.

A study of this type of system by Aroian (1) was oversimplified and had little practical value. (See review by Flehinger [23]).

Consider two extreme cases, the first of which consists of a set of n subsystems, each capable of performing the system function. At fixed time intervals T , the system output is switched from subsystem i to subsystem $i + 1$. Upon reaching subsystem n , the output is switched back to subsystem 1. It can easily be seen that the reliability of the overall redundant system is lower than the reliability of each subsystem.

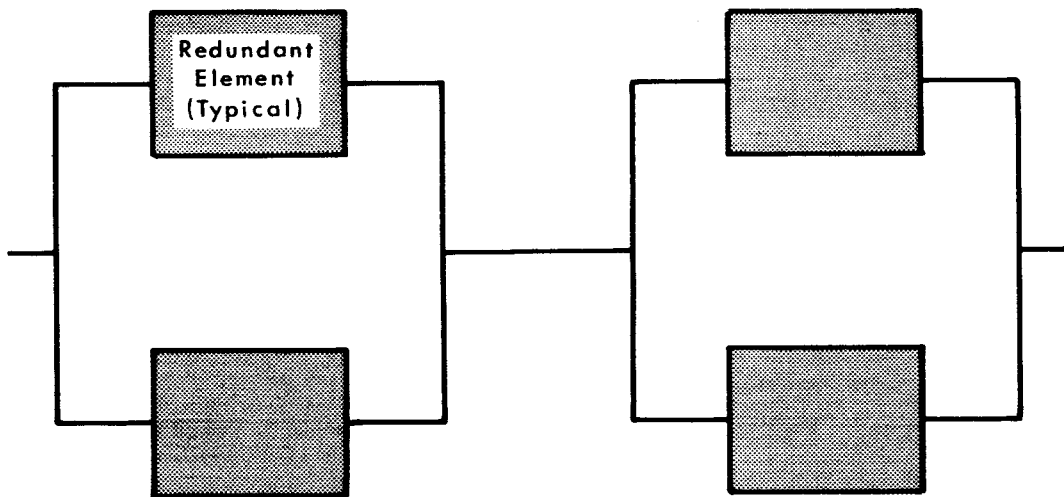
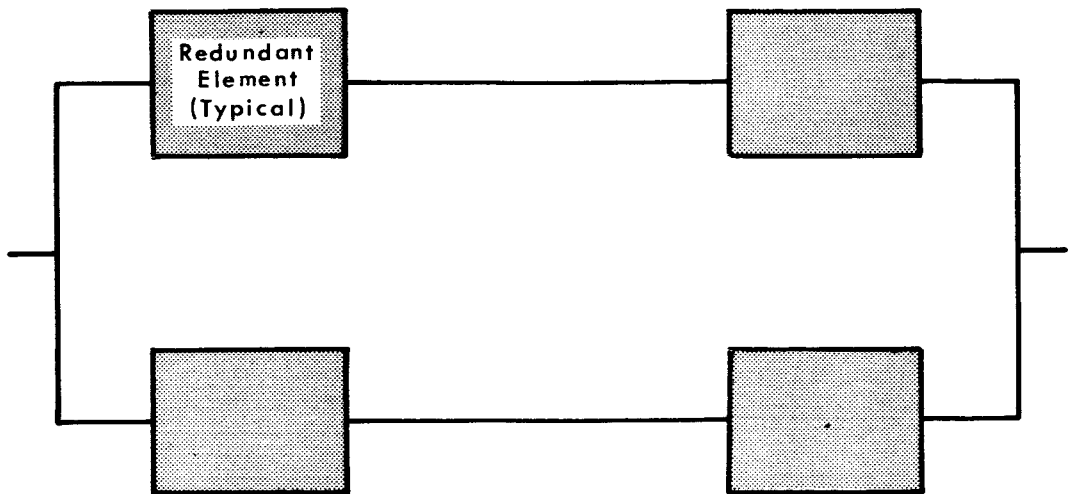


FIGURE 4. QUAD REDUNDANCY

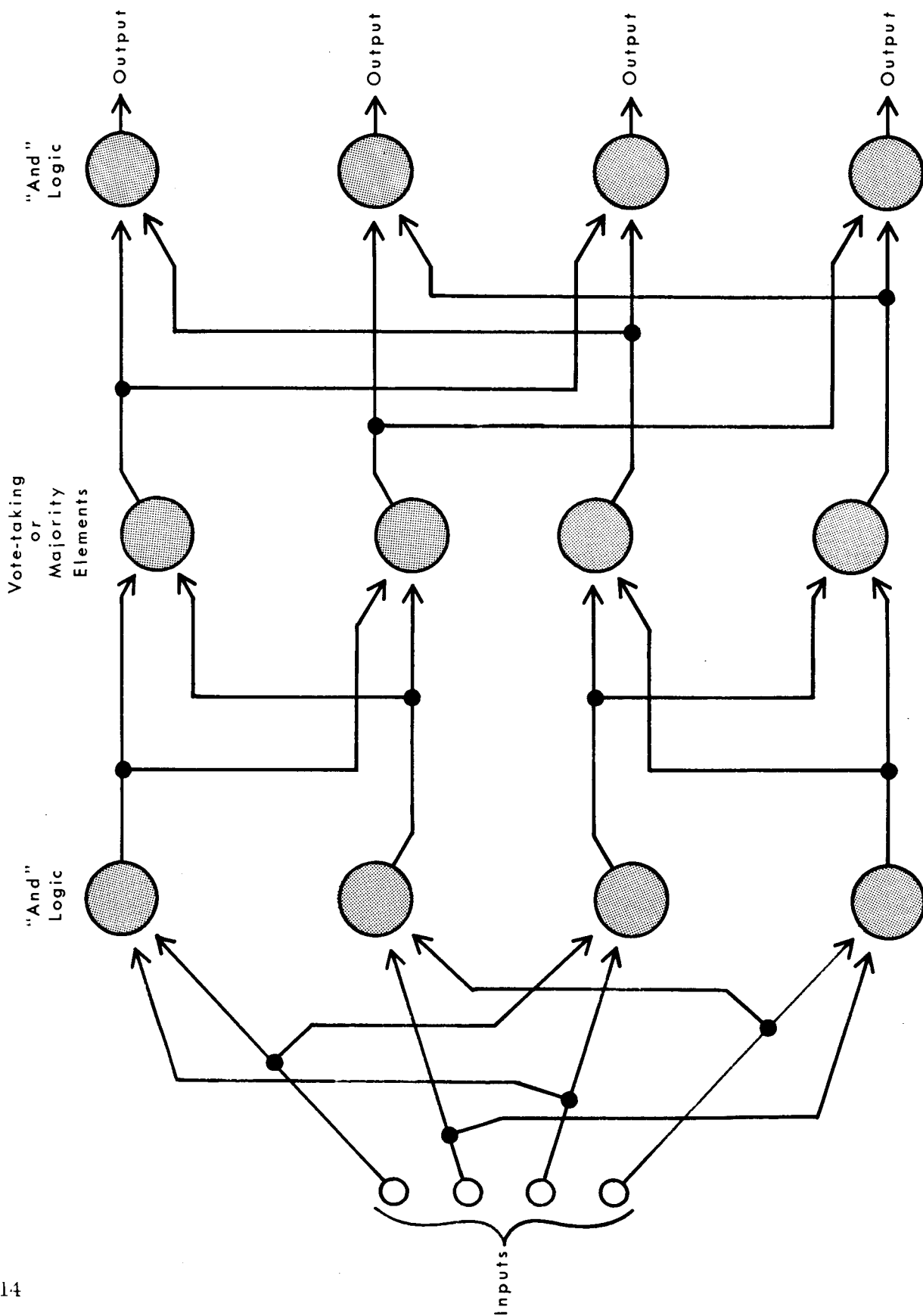


FIGURE 5. TRYON'S QUADED LOGIC

If the reliability of a subsystem is

$$(8) \quad R_s(t) = e^{-\lambda t}$$

where λ is the failure rate; then, for the period $T < t < 2T$, the system reliability is given by

$$(9) \quad R(t) = e^{-\lambda T} e^{-\lambda t} < e^{-\lambda t}$$

where $R(t)$ is the probability that the first subsystem is working at T and the second subsystem is working at t . Thus, it is better to leave the performance of system function to one single subsystem.

Consider for the second extreme case a set of n subsystems each capable of performing the useful function. One of these is active. A perfect failure detector is used for monitoring. When the active system fails, a perfect switch moves the output to the next subsystem. For this system, in the intermittent failure model, the system reliability is given by

$$(10) \quad R(t) = 1 - [1 - R_s(t)]^n$$

where R_s is subsystem reliability. $R(t)$ may thus be brought arbitrarily close to 1 by making n sufficiently large. The system mean life can be calculated for the exponential case as

$$(11) \quad M = \sum_{k=1}^n \frac{1}{k \lambda} = M_s \sum_{k=1}^n \frac{1}{k}$$

Since the sum is a divergent (infinite) series, the system mean life increases as the logarithm of the number of subsystems and the procedure is thus extremely uneconomical. Various generalizations of this type of a binomial system have been given by Kletsy (35).

An exhaustive study of the general type of system with a catastrophic failure model was reported by Flehinger (21) and remains a classic to this day. The system was divided into a variable number of modules and the described type of redundancy was applied. A failure detector and a stepper switch to change modules were included. Assumptions included that of unreliable switching and/or failure detection. Detailed formulae and curves were provided for all cases. The general conclusions were:

1. For initially reliable systems, the degree of redundancy is much more significant than the level at which it is applied.
2. If switching is imperfect, improvements are obtainable only at enormous cost.

Flehinger's general conclusions are still accepted as valid despite the hundreds of papers that have appeared since.

Redundancy with Maintenance

Maintenance offers the only economical means of significantly increasing the mean life of a system (say, by one order of magnitude). In this section maintained (and therefore manned) systems are discussed. Several new factors have to be considered:

1. The Mean Time Between Failures (MTBF), which is equivalent to the mean life of an unmaintained system if there are no subsystem failure indicators.
2. The availability of the system, which is the average fraction of time the system is operative.
3. The maintainability of the system, or the probability that a failed system is restored to operation within a specified time. Maintainability is a function of the design, maintenance procedures and logistics.

Welker and Horn (57) have discussed these factors in detail. Only catastrophic failures will be considered here.

Wired-in or active redundancy is not quite compatible with maintainability. The more redundant a system, the less maintainable it is. Indeed there are those who question the value of active redundancy in maintained systems.

A discussion of the tradeoff between redundancy and checkout has been given by Ihrig (29). There is general agreement that redundant maintained systems must be periodically inspected (checked out) to restore the redundancy (replace failed parts that have not yet resulted in system failure). It is much more difficult to detect failures in a redundant system unless individual failure detectors are included.

Johnson and Brule (33) have made a detailed study of the effect of different maintenance procedures on the reliability of a redundant system. The general conclusions are as follows,

The increase in mean-time-to-first-failure is only modest if the periodic

maintenance procedure is to examine only one module and to replace it if found defective. A redundancy ratio of 10 gives an increase of 3. However, the increase in mean-time-to-first-failure may be between 10 and 100 if all defective parts are replaced. Assumptions are that failure detectors are perfect and that the time between inspections and the repair time are both short relative to the mean life of the corresponding nonredundant system.

Rosenheim and Ash (48) have considered a related (and in some ways more practical) system. They consider redundancy on the system level, with switching and repair being done by technicians (hence perfectly). They also associate a cost with system downtime. Under these conditions, they calculate the number of machines required to optimize the system cost. Their general conclusion is that two machines are much better than one, but more than two are not justified.

Less conclusive studies of this general type have been reported by Chin (8), James et al (30) and Weisberg and Chin (56). All these studies assume an infinite supply of spare parts. If the maintenance model is changed, conclusions vary accordingly.

One obvious variation is the study of the availability of a system which is repaired at breakdown only. For a nonredundant system it is relatively simple to calculate the average repair time (57). For a redundant system, however, no one has come up with reasonable estimates of repair time, where repair is taken to mean that all the components are in perfect working order. Since availability (see Equation 2) depends on repair time, it has not been possible to establish that redundancy of the wired-in type increased availability. For redundancy at the system level (i.e., with standby systems as in the Rosenheim-Ash [48] model) we have

$$(12) \quad \bar{A}_0 = M_0 / (M_0 + T_0)$$

for the nonredundant system, and

$$(13) \quad \bar{A} = M_0 \left(\sum_{i=1}^n \frac{1}{k} \right) / \left[M_0 \left(\sum_{i=1}^n \frac{1}{k} \right) + n T_0 \right]$$

for the n-fold redundant system.

Assumptions here are that:

1. The system is operated until it breaks down.
2. All machines are repaired immediately.

If $T_0 = 0.1M_0$, we have:

$$\begin{array}{ll}\text{for } n = 2: & \bar{A} = 0.98 A_0 \\ \text{for } n = 3: & \bar{A} = 0.95 A_0 \\ \text{for } n = 10: & \bar{A} = 0.83 A_0.\end{array}$$

If in the standby system model each machine is repaired as soon as it fails (while a standby is put into service without break of service), we are back to the model studied by Johnson and Brule (33), i.e., the availability is essentially 1.0, and the mean-time-between-failures is about $100 M_0$ for a mean repair time of $0.1 M_0$. (The redundancy ratio n is not too significant as long as $n \geq 3$.)

In a variant of the Rosenheim-Ash system a finite number of spare parts is considered. One may either consider a limited stock (closed system) or a uniform or random addition to the stock of spares. Such models have been studied by Kletsky (35).

Self-repairing Redundant Systems

A logical extension of the concept of redundancy with maintenance is to incorporate the maintenance within the system. This is the concept of the self-repairing system.

Cluley (9) has discussed an extreme case of such a system. He considered a digital computer with low-level redundancy which operates in the following fashion. During short periods of interruption in real time processing, the computer goes into a diagnostic mode. By suitable switching, it is made non-redundant during this phase. Each of the redundant parts is inserted in the system and checked out. A mean life of $250 M_0$ (M_0 = mean life of nonredundant system) is obtained with a redundancy ratio of 2. These numbers immediately suggest unreasonable assumptions:

1. The portion of the computer used to execute the diagnostic program is perfectly reliable.
2. The switching necessary to improve each of the redundant parts can be done perfectly reliably in times of the order of 10^{-12} second.
3. It takes zero time for the computer to put out failure messages, and it takes zero time to replace the failed parts.
4. The diagnostic program is "ideal"; that is, it identifies the failed component.
5. The diagnostic program takes only 1 millisecond or thereabouts.

These assumptions are completely unreasonable. In current systems more than half the computer logic must be operable if diagnostic programs are to run. Clearly, this immediately limits MTBF. Since the redundancy used by Cluley is low-level, the switching required is very complicated and hence unreliable. No currently available diagnostic program comes anywhere near satisfying Assumptions 4 and 5 above. It may also be noted that Cluley's conclusions are at variance with those of Flehinger (21) even when allowing for the difference in models. Cluley claims that low-level redundancy is far superior (by a factor of 100) to system redundancy, whereas Flehinger concludes that for initially reliable machines, the level of redundancy is of secondary importance.

The most realistic study of self-repairing systems was done by Kruus (36). (Refer to Figure 6.) Kruus made the following observations:

1. The diagnosis is not perfect.
2. It takes time to repair the system.
3. It takes time to condition a digital system (load memory) before it can be placed in operation.
4. The number of repairs is reasonable (1,000) so that the switching may be done reliably.
5. Spare modules have a "shelf-life" of the same order of magnitude as the mean life of a module in operation.

The only unreasonable assumption was that the majority decision element with a dissenting vote indicator was perfectly reliable. Such a system is too complicated for analytical evaluation. The mean life was therefore obtained by computer simulation, using Monte Carlo techniques.

Kruus obtained mean lives of the order of 10 to 30 times the mean life of a simple system. The redundancy ratio was 20. The most critical variables were found to be the repair time, the conditioning time and the stress factor of spares. All parameters were assigned a given mean and distribution and were chosen randomly for the computer simulation.

The results of Kruus are far more reasonable than those of Cluley.

Commercial Applications

In this section some commercial applications of redundancy techniques are reviewed.

In the bombing navigation system of the B-58 bomber, model AN/ASQ-42V, redundancy is provided by paralleling complete systems with identical spare parts, by paralleling selected units with redundant replacements, and by providing spare units to replace any of several units with like uses. Self-sensing,

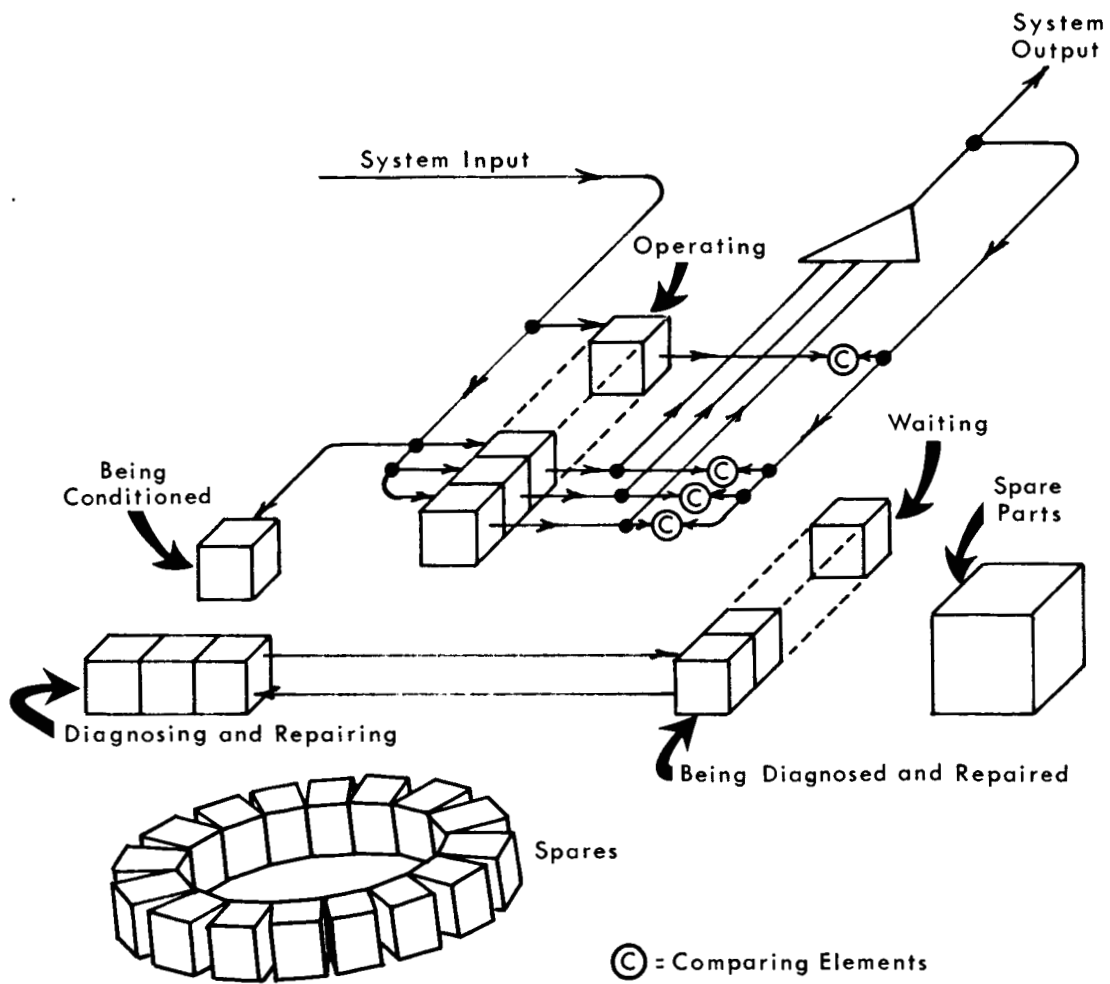


FIGURE 6. SELF-REPAIRING SYSTEM - INFORMATION TRANSMISSION

fail-safe design techniques are used. Formulae for redundant reliability are derived from the binomial distribution.

The National Aeronautics and Space Administration has set a standard of reliability of 0.987 for the modified Titan II chosen to launch the manned Gemini flights (59). There was no redundancy in the ICBM Titan, but Martin engineers have relied heavily on redundancy to improve reliability in the Gemini Titan. These redundancy provisions have added weight.

Martin stripped nonessential equipment from the Gemini Titan II and saved nearly 1,000 pounds. Vernier and retro rockets were needed for the ICBM version but not for Gemini. Titan II carried four telemetry systems as an ICBM; only two were needed for Gemini. This saved about 100 pounds.

In the ICBM version the systems used two-conductor shield wiring. This was replaced by single-conductor wiring to save almost 250 pounds.

The inertial guidance system in the ICBM model was replaced by a radio guidance system, a weight saving of about 150 pounds. The ICBM version also carried two range safety equipment systems. One of these was eliminated for the Gemini model.

However, Martin has made almost all the electronic equipment in the booster redundant to increase reliability. This has meant the addition of some fully redundant packages.

Martin has developed a new malfunction detection system to give astronauts the safety features needed for manned space flight. This system monitors every critical area of the launch system and immediately flashes a warning when trouble is indicated. Complete redundancy is provided.

General Electric Company's Light Military Electronics Department (11 and 26) has put into production a redundant Digitizer Logic Unit designed and developed for use in the OAO (Orbiting Astronomical Observatory), which is scheduled for launch by the National Aeronautics and Space Administration in 1965.

The DLU (Digitizer Logic Unit), part of the stabilization and control system of the satellite, generates dc analog error signals for correction of attitude and tracking errors. This unit takes inputs from six star trackers (each containing two gimbals), compares these inputs with command information, computes the error, and puts out an analog signal proportional to the error. An analog signal is then provided which is proportional to the error for each star tracker gimbal in a time multiplex system.

The analog section error output is from -10 to + 10 volts and is presented in 39-millivolt steps. Matched-pair transistors, full temperature compensation and initial trimming to within 1 mv from the dc amplifier outputs were used to achieve the specified high accuracy over the full range of environments. In some of the redundant switches, transistor parameters such as inverse beta had also to meet stringent tolerances.

Reliability requirements dictated the design of a redundant system. These reliability requirements were 0.98 for 12,000 hours allowing 1 out of 12 gimbal failures during this time. This period includes 1 year of orbital life plus total shelf life.

General Electric designed a redundant system based primarily on digital function triplication and majority voting to meet these requirements. The analog section also uses component redundancy wherever possible. Component count is in excess of 8,000.

To allow full exploitation of the inherent redundant system design, it was shown that circuit and component redundancies required redundant interconnection techniques. As a result the following approach was used.

Any function signal that is derived only once has at least two paths from its origin to all destination points. Similarly, all power and ground lines have two or more paths to all connection points. Reliance on a single mechanical joint, i.e., solder or weld, is circumvented by the use of two or more electrically identical but physically separate origin and destination points. Those signals that are derived independently in triplicate are handled on single interconnecting lines since the interconnection is not the weak point in the chain, as would be the case for the singly derived function. Connector/connector-interfaces utilize two separate pins per signal or power function.

The interconnections present in the DLU thus provide redundant signal flow paths which augment the redundant functional and circuit designs.

An automated continuity test is employed on the main wiring harness. This test checks over 1,200 wires.

The main points of concern are the origin, the destination, and the solder joints. At present, the solder test is one of individual inspection. Work is under way on infrared detection techniques to eliminate the human factor.

While components are not the weak point as far as reliability is concerned, there are still a number of problems involved in component testing in a redundant system, especially when feedback loops are encountered.

In a triplicated majority vote system with feedback, the inputs in feedback signals for a given channel can be correct even though a failure exists within the channel, since the redundancy employed can mask the internal failure. Thus, it is necessary to exercise the entire loop, circuit by circuit, and almost component by component.

It is predicted that even with the advent of integrated circuit techniques and hardware due in the next 5 to 10 years, redundancy will still remain the answer to more and more stringent customer requirements.

GE's technique for use with digital circuits utilizes three parallel channels, any one of which could carry on the operation if redundancy were not wanted. The three parallel channels are interconnected at suitable junctures allowing a majority vote to be taken to determine the presence or absence of pulses at any instant. If a malfunction occurs in any one channel, the other two channels will carry the "vote" to produce the correct result.

The system may experience a sizable number of faults in all three channels and still work satisfactorily if the faults do not occur at the same time in identical segments of two out of three channels. This may be made possible by dividing each of the parallel channels into a large number of segments and taking a majority vote after each segment.

"Level of voting" is the term applied to the number of such segments or points of vote-taking.

A three-channel redundant system with only a single level of voting should be able to operate for a period four times longer than the period for a non-redundant system with a 98 percent probability of not failing. For the same system with a 20-level of voting, the operating period would jump to 17 times the operating period of a nonredundant system with a 98 percent probability.

Five or seven parallel channels could be used with a three-out-of-five or a four-out-of-seven majority vote. The increased cost and complexity would offset the small gain in reliability so that use of three parallel channels is the chosen configuration.

The use of parallel channel redundancy with many levels of voting increases the number of components required by a factor of at least three with a lesser increase in overall equipment size and weight. Advances in microelectronics make this technique adaptable for many defense and space uses.

GE has built a redundant logic keystream generator to show what can be done with the present state of the art. This generator is fabricated with thin-film techniques and with the use of bonded pico transistors. This redundant

(three channel, five-stage) circuit contains 116 transistors and occupies a volume of 1 cubic inch. The logic circuits operate at a 5 mc clock rate and the entire generator consumes less than 250 milliwatts.

Achievement of Reliability

Skilled product design and production is the best guarantee of reliability. Electronic and aeronautical engineers work within many restrictions which affect their designs. Some of these restrictions are imposed by cost, weight, volume, and configuration. Two other considerations are that the ultimate in design is seldom attainable, and that relevant data are not always at hand.

Redundancy is a technique which employs more than one method to bring about increased reliability. The redundancy concept is not new. Dual wheels on trucks and airplane landing gears are early examples of redundancy.

Maintenance is also important to reliability. Failures may be avoided by testing or checking certain elements and replacing defective items or parts. For this reason it is essential that parts be replaceable and interchangeable.

Built-in test equipment is one method of securing reliability. Electronic technicians generally use simple equipment of the go-no-go and press-to-test type. Methods of achieving reliability also include debugging and testing to destruction. Still another method is 100 percent testing under extreme environmental conditions.

There are many manufacturing practices which affect reliability: e.g., quality control, storage, and workmanship. The Department of Defense, in many of its contracts, specifies reliability as part of the requirements of the contract.

Redundancy Not a Cure-all

Redundancy is not a reliability cure-all. It complements but does not replace the need for a strong effort to ensure reliability in other technical and program areas. For example, if any one part type in a satellite is not capable of withstanding the launch vibration, the redundant use of such parts would not help since all such redundant parts would fail concurrently and satellite reliability would drop to zero.

Bazovsky (4) stated that "the lack of reliability wastes billions of dollars

and has slowed technological progress in many vital areas. There is perhaps no engineering field today where the need for improvement is greater than in the field of reliability."

REFERENCES

References 3, 7, 31, 34, 37, 41, and 52 of this reference list contain extensive bibliographies on the subject of redundancy and reliability.

1. Aroian, L. A.: "The Reliability of Items in Sequence with Sensing and Switching, " Redundancy Techniques for Computing Systems, 318-327, Spartan Books, Inc. (1962).
2. Balaban, H. S.: "Some Effects of Redundancy on System Reliability, " Proceedings of the Sixth National Symposium of Reliability and Quality Control, 388-402 (January 1960).
3. Balaban, H. S.: "A Selected Bibliography on Reliability, " ARINC Research Corporation, Publication 4553-290 (February 1962).
4. Bazovsky, I.: "Reliability Theory and Practice, " Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1961).
5. Birnbaum, Z. W., Esary, J. D., Saunders, S. C.: "Multi-Component Systems and Structures and their Reliability, " Technometrics, 3 (1) 55-77 (February 1961).
6. Brown, W. G., Tierney, J., Wasserman, R.: "Improvements of Electronic Computer Reliability Through the Use of Redundancy, " Institute of Radio Engineers, Transactions on Electronic Computers, EC-10, 407-416 (September 1961).
7. Carroll, J. M.: "Reliability 1962, " Electronics, 35, 53-76 (November 30, 1962).
8. Chin, J. H. S.: "Group Redundancy in Space Electronics, " 19th National Meeting of the Operations Research Society, Chicago (May 25, 1961).
9. Cluley, J. C.: "Low Level Redundancy as a Means of Improving Digital Computer Reliability, " Electronics Reliability and Microminiaturization, 1, 203-216, Pergamon Press (1962).

10. Cohn, M.: "Redundancy in Complex Computers," Proceedings of the National Conference on Aeronautical Electronics, Dayton, 231-235 (May 1956).
11. Costa, J.: "Redundant Digitizer Logic Unit for OAO (Orbiting Astronomical Observatory)," Electronic News, 5 (October 14, 1963).
12. Cox, D. R., Smith, W. L.: "On the Superposition of Renewal Processes," Biometrika, 41, 91-99 (1954).
13. Creveling, C. J.: "Increasing the Reliability of Electronic Equipment by the Use of Redundant Circuits," Proceedings of the Institute of Radio Engineers, 44, 509-515 (1956).
14. dePian, L.: "Reliability Using Redundancy Concepts," George Washington University, Technical Report (1959) AD-210-692.
15. dePian, L., Grisamore, N. T.: "Two Approaches to Incorporating Redundancy into Logical Design," "Redundancy Techniques for Computing Systems," 379-388, Spartan Books, Inc. (1962).
16. Dickinson, W. E., Walker, R. M.: "Reliability Improvement by the Use of Multiple-Element Switching Circuits," IBM J Res, 2, 142-147 (1958).
17. Dostart, L. J., McGill, J. A., Baechler, D. O.: "Application of von Neumann Redundancy Techniques to the Reliable Design of Digital Computers," ARINC Research Corporation, Final Technical Summary Report, Contract No. AF30(602)-2419 (April 1962).
18. Einhorn, S. J.: "Reliability Prediction for Repairable Redundant Systems," Proceedings of the Institute of Electrical and Electronic Engineers, 51, 312-317 (February 1963).
19. Farrell, E. J.: "Improving the Reliability of Digital Devices with Redundancy," Institute of Radio Engineers, Transactions on Reliability and Quality Control, RQC-11, 44-50 (May 1962).
20. Fasano, R. M., Lemak, A. G.: "A Quad Configuration - Reliability and Design Aspects," Proceedings of the Eighth National Symposium on Reliability and Quality Control, 394-407 (1962).
21. Flehinger, B. J.: "Reliability Improvement Through Redundancy at Various System Levels," IBM J Res, 2, 148-158 (April 1958).

22. Flehinger, B. J.: "Review of Farrell," Institute of Radio Engineers, Transactions on Electronic Computers, EC-11, 798 (December 1962).
23. Flehinger, B. J.: "Review of Aroian," Institute of Electrical and Electronic Engineers, Transactions on Electronic Computers, EC-12, 33 (February 1963).
24. Flehinger, B. J.: "Review of Price," Institute of Electrical and Electronic Engineers, Transactions on Electronic Computers, EC-11, 32 (February 1963).
25. Gordon, R.: "Optimum Component Redundancy for Maximum System Reliability," Operations Res, 5, 229-243 (April 1957).
26. Hall, K. L., Harbach, A. B., Herbert, C. H., LaCapra, J.: "Data Handling Equipment (OAO) Redundant Design," Paper Presented at the National Symposium on Space Electronics and Telemetry, Miami Beach (October 2-4, 1962).
27. Hall, K. L.: "Basic Rules for Designing," Electronics, 36, 62-66 (April 1963).
28. Hall, K. M., McDonald, R. H.: "Improving System Reliability," Proceedings of the Seventh National Symposium on Reliability and Quality Control, 214-228 (January 1961).
29. Ihrig, W. E.: "Reliability Tradeoff," Electronic Design, 73, 73-77 (May 1963).
30. James, D. C., Kent, A. H., Holloway, J. A.: "Redundancy and Detection of First Failures," Paper Presented at WESCON (August 1961).
31. Jensen, P. J.: "Bibliography on Redundancy Techniques," "Redundancy Techniques for Computing Systems," 389-403, Spartan Books, Inc. (1962).
32. Jervis, E. R.: "Reliability in Microelectronics," Paper Presented at the Conference on the Impact of Microelectronics, Chicago (June 27, 1963).
33. Johnson, R. A., Brule, J. D.: "Diagnosis of Equipment Failures," Syracuse University, Final Report, Contract No. AF(30)-602-1833.
34. Jorgensen, W. E., Carlson, I. G., Gross, G. G.: "NEL Reliability Bibliography," Naval Electronics Laboratory, San Diego, California (1956).

35. Kletsy, E. J.: "Self-Repairing Machines, " Part I of Final Report, Syracuse University, Contract No. AF(30)-602-2234 (1961).
36. Kruus, J.: "Upper Bounds for the Mean Life of Self-Repairing Systems, " Technical Report R-172, Coordinated Science Laboratory, University of Illinois (1963).
37. Luebbert, W. F.: "Literature Guide on Failure Control and Reliability, " Technical Report 13, Stanford University (December 1956).
38. Lyons, R. E., Vanderkulk, W.: "Use of Triple Modular Redundancy to Improve Computer Reliability, " IBM J Res, 6, 200-209 (April 1962).
39. Maitra, K. K.: "Review of dePian and Grisamore, " Institute of Electrical and Electronic Engineers, Transactions on Electronic Computers, EC-12, 32 (February 1963).
40. Moore, E. F., Shannon, C. E.: "Reliable Circuits Using Less Reliable Relays, " J Franklin Inst, 262, 191-208, 281-297 (1956).
41. Motes, J. H.: "KWIC Index to Reliability and Quality Control Literature, " Proceedings of the Ninth National Symposium on Reliability and Quality Control, 556-581 (1963).
42. Nagy, G.: "Reliability of Repairable Systems, " Proceedings of the Ninth National Symposium on Reliability and Quality Control, 93-108 (1963).
43. Pierce, W. H.: "Improving Reliability of Digital Systems by Redundancy and Adaption, " Stanford University (1960).
44. Pierce, W. H.: "Adaptive Vote-takers Improve the Use of Redundancy, " "Redundancy Techniques for Computing Systems, " 229-250, Spartan Books, Inc. (1962).
45. Pierce, W. H.: "Adaptive Decision Elements to Improve the Reliability of Redundant Systems, " Institute of Radio Engineers, National Convention Record, 10, (4) 124-131 (1962).
46. Plait, A.: "When is Reliability Improved by Quad Redundancy?" Space/Aeron, 37, 77-82 (March 1962).
47. Price, H. W.: "Mean Life of Parallel Electronic Computers - Exponential Distribution Case, " "Redundancy Techniques for Computing Systems, " 304-317, Spartan Books, Inc. (1962).

48. Rosenheim, D. E., Ash, R. B.: "Increasing Reliability by the Use of Redundant Machines," Institute of Radio Engineers, Transactions on Electronic Computers, EC-8, 125-130 (June 1959).
49. Schwartz, L. S.: "Reliability Through Redundancy and Error Correcting Codes," Electro-Technology, 67, 123-130 (February 1961).
50. Sorensen, A. A.: "Digital Circuit Reliability Through Redundancy," Electro-Technology, 68, 118-125 (1961).
51. Teoste, R.: "Design of Repairable Redundant Computers," Institute of Radio Engineers, Transactions on Electronic Computers, EC-11, 643-49 (October 1962).
52. Trew, J. R.: "Circuit Design Reliability Through Redundancy, A Literature Survey," Space Technology Laboratories Report (August 1960).
53. Tryon, J. G.: "Quadded Logic," "Redundancy Techniques for Computing Systems," 205-28, Spartan Books, Inc. (1962).
54. von Neumann, J.: "Probabilistic Logics," "Automata Studies," Princeton University Press, 43-98 (1956).
55. Weinstock, G. D.: "Matrix Analysis of Reliability for One-Shot Redundant Systems," Electro-Technology, 70, 99-103 (November 1962).
56. Weisberg, S. A., Chin, J. H. S.: "Reliability and Availability of Some Redundant Systems," Paper Presented at the Maintainability Conference, Bedford, Massachusetts (March 12-13, 1963).
57. Welker, E. L., Horne, R. C.: "Concepts Associated with System Effectiveness," ARINC Research Corporation, Monograph No. 9, Publication 123-4-163 (July 1960).
58. Wilcox, R. H., Mann, W. C.: "Redundancy Techniques for Computing Systems," Spartan Books, Inc., Washington D. C. (1962).
59. "Launch Vehicles," Product Engineering, 34, 64 (October 28, 1963).